

Network-based model of the growth of termite nests

Young-Ho Eom¹, Andrea Perna², Santo Fortunato^{3,4}, Eric Darrouzet⁵, Guy Theraulaz^{6,7}, Christian Jost^{6,7,*}

¹*IMT Institute for Advanced Studies Lucca, Piazza San Francesco 19, Lucca 55100, Italy*

²*Laboratoire Interdisciplinaire des Energies de Demain - Paris Interdisciplinary Energy Research Institute, Paris Diderot University, 10 rue Alice Domon et Léonie Duquet, Paris, France*

³*Complex Systems Unit, Aalto University School of Science, P.O. Box 12200, FI-00076, Finland*

⁴*Center for Complex Networks and Systems Research, School of Informatics and Computing, Indiana University, Bloomington, Indiana 47405 USA* ⁵*IRBI, UMR CNRS 7261, University of Tours, Faculty of Sciences, Parc de Grandmont, 37200 Tours, France*

⁶*Université de Toulouse, UPS, CRCA (Centre de Recherches sur la Cognition Animale), Bât 4R3, 118 route de Narbonne, F-31062 Toulouse, France*

⁷*CNRS, CRCA, Bât 4R3 118 route de Narbonne, F-31062 Toulouse, France*

**Corresponding author: CJ (christian.jost@univ-tlse3.fr)*

(Dated: December 11, 2015)

We present a model for the growth of the transportation network inside nests of the social insect subfamily Termitinae (Isoptera, termitidae). These nests consist of large chambers (nodes) connected by tunnels (edges). The model based on the empirical analysis of the real nest networks combined with pruning (edge removal, either random or weighted by betweenness centrality) and a memory effect (preferential growth from the latest added chambers) successfully predicts emergent nest properties (degree distribution, size of the largest connected component, average path lengths, backbone link ratios, and local graph redundancy). The two pruning alternatives can be associated with different genuses in the subfamily. A sensitivity analysis on the pruning and memory parameters indicates that Termitinae networks favor fast internal transportation over efficient defense strategies against ant predators. Our results provide an example of how complex network organization and efficient network properties can be generated from simple building rules based on local interactions and contribute to our understanding of the mechanisms that come into play for the formation of termite networks and of biological transportation networks in general.

PACS numbers: 89.75.Hc, 89.75.Fb, 05.65.+b, 87.23.Cc

I. INTRODUCTION

Biological transportation networks are a fundamental component of all living systems, allowing the exchange of information and material at the scale of the whole system. It is unsurprising that structures specialized for transportation are found at all levels of biological organization, including intracellular transportation along the cytoskeleton, and vascular and neural networks within the body of animals. At the biological scale of animal groups and communities, the transportation of food and material, the movements of animals and the inter-individual encounters are often supported by specialized transportation networks of trails, galleries, and burrows that the animals produce and use in their exploration and foraging movements (reviewed in [1]). Social insects in particular are known to produce some of the most complex networks of trails and galleries in the animal kingdom, probably as a result of their high level of sociality. These networks include trail networks in ants or termites [2–4], and underground systems of tunnels formed by ants [5] or termites [6–8].

All biological transportation networks share similar functions: by favoring transportation over distances much larger than those permitted by simple diffusion they mediate the integration of the different parts that compose a biological system, supporting the functional

unity of the system as a whole. They also share a similar morphogenesis, in the sense that almost all biological transportation networks are produced as the result of self-organized (SO) morphogenetic processes whereby the growth is driven by locally available information, in the absence of a pre-existing master plan of the network [9, 10]. The formation of animal and human transportation networks has been modeled with models based on growth alone (e.g., army ant raid networks [11], ant gallery networks [5, 12], or urban street patterns [13]), pruning of an existing network [14, 15], or a combination of both [16–18].

A common characteristic of all these networks is that they are embedded in a 2D or 3D environment, that is, both the position of network nodes and the layout of network edges are associated with sets of spatial coordinates. The effects of spatial embedding cannot be neglected when trying to understand the formation and topological properties of these networks, because the probability of existence of a connection between two nodes depends not only on their relative distance [19], but also on the physical arrangement and steric interactions between edges. While a large part of the existing network literature has dealt with social or communication networks, which are comparatively less affected by spatial constraints (see reviews in [20–22]), the theoretical foundations underlying the analysis and the modeling of spatially explicit net-

works are comparatively less developed (see [23] for a review) and often deal with specific fields such as urban transportation and human mobility patterns [24].

In this paper we focus on a specific class of biological transportation networks represented by the network of chambers and tunnels that termites of the subfamily Termitinae produce in the above-ground part of their nest. We identify how structural features of these transportation networks emerge through self-organization based on local rules. The nodes of these networks correspond to chambers and the edges to tunnels connecting these chambers. These networks provide safe living space to termites and are connected to an underground tunnel network (not considered in this study) that connect the nest to foraging grounds [25–27]. Termite movements along these networks involve both bringing back food and distributing it to the colony (e.g. termite soldiers, larvae and the royal couple) and daily patterns of movements of various individuals to different parts of the nest with favorable environmental conditions.

By analyzing an extended dataset containing nests from three Termitinae genera: *Cubitermes*, *Proculitermes* and *Thoracotermes*, we formulate a model for the growth of the transportation networks internal to the nests of these different termites. The model has two free parameters, λ that controls preferential growth on the periphery and ξ that determines the pruning or removal of existing edges. Two variants of pruning, random pruning or pruning weighted by betweenness centrality, are explored, and we compare both variants to random geometric graphs. We calibrate λ and ξ to the real nests and show that network properties not used in the calibration process are faithfully reproduced by the models with pruning, the variant based on betweenness centrality performing slightly better. The validated models are then used to assess the sensitivity of some network properties, linked to internal transportation or nest defense, to the two free parameters in order to interpret the values that have been found for the real nests.

II. MATERIALS AND METHODS: FROM AN EMPIRICAL NEST ANALYSIS TO THE NEST GROWTH MODEL

A. The available termite nests

The 12 networks modeled in this paper (see Table I) correspond to the above ground part of termite nests from the African continent. They all belong to the Termitinae sub-family that have a common architecture: distinct chambers with a large diameter interconnected by tunnels of a small diameter (these different elements can be easily identified because in these nests the connecting tunnels have a diameter about 10 times smaller than the chambers' diameter). In our network representation, chamber barycentres are associated to nodes and tunnels between chambers to edges between the corresponding

nodes. The detailed network extraction method from x-ray tomographies of the nests has been described in [28]. Node coordinates are converted from voxel positions to metric (x, y, z) coordinates. The z -axis corresponds to the vertical direction.

The majority of the nests used in our study have been obtained from natural history museums in France, with their taxonomic identity only known to the genus level. There are six *Cubitermes* nest networks (already published in [28]), four *Proculitermes sjostedti* networks collected in 2007 in Côte d'Ivoire, and two *Thoracotermes* networks (the larger one, a *T. macrothorax*, was collected in the Republic of the Congo in 2009). See the SM [29] for the general shape of all nests. While some aspects of mound architecture are typical of nests of each genus (mushroom like shapes in *Cubitermes*, straight pillars in *Thoracotermes*, and conic forms in *Proculitermes*), attempts to use this architecture for taxonomy have failed [30], pointing to the fact that all the nests in the Termitinae subfamily (hence all the nests analyzed here) share similar morphological characteristics and possibly result from similar construction rules.

B. Empirical network analysis

Each tunnel or edge in the nest networks can be considered as a three dimensional vector represented by Cartesian coordinates (x, y, z) or by spherical coordinates (r, θ, ϕ) . The latter will be used to characterize the nests. Since there is no natural orientation in a tunnel connecting chambers $\vec{c}_1 = (x_1, y_1, z_1)$ and $\vec{c}_2 = (x_2, y_2, z_2)$, it can be represented as either $\vec{c}_1 - \vec{c}_2$ or as $\vec{c}_2 - \vec{c}_1$ (Fig. 1(a)). Both vectors are used in the empirical description of the nests, e.g., Fig.1(b-d) for nest C9 that shows the distributions of tunnel length r , the vertical component θ and the horizontal component ϕ . The distributions for the other 11 nests are shown in the SM [29].

Both the r distribution (Fig. 1(b)) and the θ distribution (Fig. 1(c)) resemble normal distributions, the latter with mean $\pi/2$. However, the distribution of ϕ ($x - y$ plane) rather follows a uniform distribution (Fig. 1(d)). These patterns are confirmed in the other nests (see SM [29]). This means that each termite nest can be characterized by three parameters: the mean and standard deviation of its tunnel length distribution $(\bar{r}, \bar{\sigma}_r)$, and the standard deviation of the vertical θ direction component $(\bar{\sigma}_\theta)$.

C. Model description

Based on the above empirical observations we propose a simple nest growth model, betweenness based pruning (BBP) model, with five main procedures: (i) set nest boundaries and initial node, (ii) determine the initiation node for the next tunnel, (iii) construct an edge (iv) cre-

TABLE I. General statistics of the analyzed termite nest networks and the model parameters that have been calibrated to each nest (see model description). Nr-C: number of chambers; Nr-T: number of tunnels; ND: average node degree. LCC: size of the largest connected component. $(\lambda_{BBP}, \xi_{BBP})$ are the parameters for the BBP model, (λ_{RP}, ξ_{RP}) the ones for the RP model, and R_{RGG} is the parameter for the RGG model.

Nest	Genus	Nr-C	Nr-T	ND	LCC	λ_{BBP}	ξ_{BBP}	λ_{RP}	ξ_{RP}	R_{RGG}
C9	<i>Cubitermes</i>	532	682	2.56	507	0.021	0.117	0.020	0.126	0.199
C10	<i>Cubitermes</i>	396	371	1.87	349	0.065	0.292	0.065	0.292	0.197
C11	<i>Cubitermes</i>	344	310	1.80	260	0.066	0.279	0.065	0.274	0.205
C12	<i>Cubitermes</i>	190	234	2.46	183	0.067	0.246	0.066	0.244	0.274
C18	<i>Cubitermes</i>	312	343	2.2	287	0.031	0.250	0.029	0.248	0.226
C19	<i>Cubitermes</i>	295	445	3.02	268	0.045	0.085	0.046	0.080	0.258
P67	<i>Procubitermes</i>	1123	2149	3.83	1091	0.025	0.0	0.025	0.0	0.177
P78	<i>Procubitermes</i>	675	878	2.60	598	0.014	0.046	0.014	0.044	0.185
P79a	<i>Procubitermes</i>	440	525	2.39	347	0.006	0.290	0.006	0.289	0.208
P79b	<i>Procubitermes</i>	388	383	1.97	292	0.0	0.451	0.0	0.447	0.203
T29	<i>Thoracotermes</i>	98	96	1.96	90	0.033	0.551	0.033	0.515	0.303
T82	<i>Thoracotermes</i>	1073	1470	2.74	1069	0.018	0.177	0.018	0.164	0.159
mean						0.033	0.232	0.032	0.227	0.216

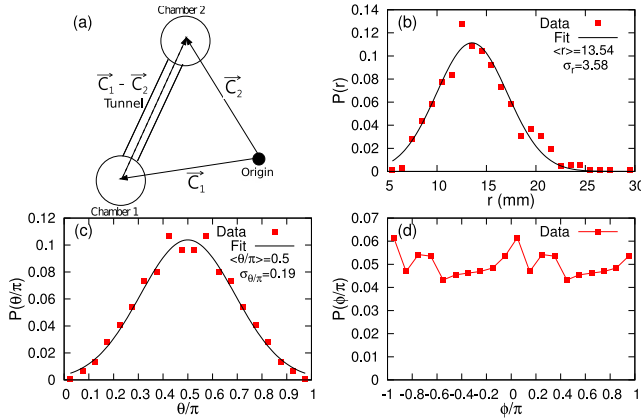


FIG. 1. (Color online) Empirical network description in the case of nest C9 (*Cubitermes* sp., Central African Republic, see Fig. 1B in [28]). (a) Each tunnel or edge is represented as a vector in spherical coordinates, (b) tunnel lengths (r) distribution, (c) distribution of the vertical component (θ , 0 points upwards and π points downwards), and (d) distribution of the horizontal component (ϕ). (c) and (d) are symmetric and periodic respectively because both $\vec{c}_1 - \vec{c}_2$ and $\vec{c}_2 - \vec{c}_1$ are used.

ate a new node (and give it an increasing unique identification number $i \geq 1$) or connect to an existing node if it is close to the end of the new edge, (v) prune edges of lesser importance.

- (i) **Initial and boundary condition:** From the empirical nest data we can compute the x-y-z intervals $[x_{min}, x_{max}]$, $[y_{min}, y_{max}]$ and $[z_{min}, z_{max}]$ for each nest. We impose initial and boundary conditions based on these intervals. We then choose an initial node (x_0, y_0, z_0) randomly in the intervals $x_0 \in [0.75x_{min} + 0.25x_{max}, 0.25x_{min} + 0.75x_{max}]$,

$y_0 \in [0.75y_{min} + 0.25y_{max}, 0.25y_{min} + 0.75y_{max}]$ and $z = z_{min}$. During nest growth we assume boundary conditions which have the shape of an ellipse in the $x - y$ plane. The formula of the ellipse is given by

$$\left(\frac{x - X_c}{A}\right)^2 + \left(\frac{y - Y_c}{B}\right)^2 = 1 \quad (1)$$

with $X_c = (x_{min} + x_{max})/2$, $Y_c = (y_{min} + y_{max})/2$, $A = (x_{max} - x_{min})/2$ and $B = (y_{max} - y_{min})/2$. The z coordinate cannot be below z_{min} , but there is no upper limit for z . The biological rationale behind these choices is that termite nest volumes are proportional to colony size [31, colony size is the number of individuals living in a single colony] and that nests are not enlarged during colony growth but rather rebuilt from scratch - colony size can therefore provide a template for the surface of the initial nest site construction and for the final nest height.

- (ii) **Determining the node of the next edge's origin:** At each time step we choose a node randomly from all existing ones. However, in order to foster nest expansion (growth from peripheral nodes) we give preference to the latest added nodes by choosing node i with probability $\exp(-\lambda(n - i))$. Here n is the total number of nodes for the given nest. The characteristic time constant λ is specific to each nest and will be chosen by calibration (see below). The biological idea is that termites mark construction material with some volatile chemical [32] and prefer recently built elements to continue construction.
- (iii) **Direction and length of the next edge:** The length (r) and the directions (θ , ϕ) of the next edge are chosen randomly from the normal or uniform distributions determined above. Edge length

r is restricted to $r \geq r_{min}$ where r_{min} is the observed minimum edge length (see also the next rule, shorter edges would lead to an edge connection to the original node). The vertical component θ is restricted to the interval $(0, \pi)$. Furthermore, since nest construction goes upward and since edges in the original data are not oriented we avoid excessive downward construction by replacing a $\theta > \pi/2 + \sigma_\theta$ by $\theta = \pi - \theta$ (recall that $\theta = \pi$ points downwards). Finally, ϕ is drawn randomly from a uniform distribution in $(0, 2\pi)$.

- (iv) **Edge construction:** This new edge is only added if it does not quit the ellipsoid boundary condition (otherwise a new edge (r, θ, ϕ) is drawn until it remains within these boundaries). The biological rationale is that termites sense gravity [33] and do not extend construction over empty space (note that the construction of the characteristic “hats” in *Cubitermes* nests is not included in our model). A new node is created if there is no existing node within the distance r_{min} of the endpoint of the new edge. If there already exists a node within this distance the new edge is connected to this node.
- (v) **Pruning of edges:** At each time step i we compute the edge-betweenness of each edge [34], that is the number of shortest paths between pairs of nodes that pass through this edge, and remove the one with the smallest value with probability ξ . Such pruning is a common feature of most observed transportation networks (reviewed in [1]). It has been directly observed in termite’s underground tunneling networks [8, 35], and there is an indirect evidence that it happens in *Cubitermes* nest growth [28].

Repeat steps (ii-v) until the model has the same number of nodes as the original nest. The biological rationale underlying this stop criterion is again the observation that mound size is proportional to colony size [31] and mound size is better approximated by the number of chambers (volume) than the number of edges (network length).

Since λ is related to nest height h and ξ is related to the total number of tunnels L , we determine these two parameters one by one, first estimating λ (with an arbitrarily fixed ξ) by minimizing the error function

$$\epsilon_1 = |h_{data} - h_{model}|/h_{data} \quad (2)$$

with a standard bisection method, then estimating ξ (fixing the already estimated λ) by minimizing the error function

$$\epsilon_2 = |L_{data} - L_{model}|/L_{data}. \quad (3)$$

For each value of (λ, ξ) we simulate 500 nests in order to compute mean h_{model} and L_{model} . Overall error is defined as $\epsilon = \epsilon_1 + \epsilon_2$. We will use an analysis of variance

to test whether the estimated parameters are specific for a genus, reporting the resulting F statistic with the associated degrees of freedom and p -values.

D. Alternative models: Random geometric graph (RGG) model and random pruning (RP) model

We consider two alternative models to clarify the performance of the BBP model and the roles of local rules in the BBP model: (i) random geometric graph model and (ii) random pruning model. A random geometric graph (RGG) is the mathematically simplest spatial network [23]. The RGG model for a given nest is embedded in the cylindrical space having the same interval of $[x_{min}, x_{max}]$, $[y_{min}, y_{max}]$, and $[z_{min}, z_{max}]$ obtained from the nest with the ellipsoid boundary condition given by Eq. (1). The nodes of the RGG model are uniformly distributed in this cylinder. Two nodes i and j are connected if the below condition is satisfied:

$$\left(\frac{X_i - X_j}{A}\right)^2 + \left(\frac{Y_i - Y_j}{B}\right)^2 + \left(\frac{Z_i - Z_j}{C}\right)^2 \leq R_{RGG}^2 \quad (4)$$

Here X_i , Y_i , and Z_i indicate x , y , and z coordinates of node i , respectively and $A = (x_{max} - x_{min})/2$, $B = (y_{max} - y_{min})/2$, and $C = (z_{max} - z_{min})/2$. We determine the value of R_{RGG} such that it gives us the same number of links in the RGG network as in the original nest network.

The random pruning (RP) model is the same as the original model in every points (i.e., from (i) to (iv) in the previous subsection) except the pruning process (i.e., (v) in the previous subsection) where we remove a randomly selected link with probability ξ rather than the lowest edge-betweenness link.

E. Model validation

We will use five emergent properties to compare between the original and the simulated networks: (i) node degree distribution, (ii) size of the largest connected component, (iii) average topological path length in the largest connected component, (iv) backbone link ratio (fraction of edges whose removal leads to a disconnection of the largest connected component), and (v) local graph redundancy (as defined in [36], it complements backbone link ratio by computing the mean of the inverse of the topological path length to connect two adjacent nodes once the direct link has been blocked: low values indicate long detours). Predicted properties will be based on 1000 simulated networks.

F. Sensitivity analysis

We assess the influence of λ , ξ , and network size (number of nodes) on four network properties: (i) size of the

largest connected component, (ii) average topological distance between any two nodes of the largest connected network component, (iii) backbone link ratio, and (iv) local graph redundancy. For an efficient transport inside the nest and easy defense against predators (such as ants) termite nests should show low average distance (fast transport), a high backbone link ratio (tree-like structures, tunnel blocking by a soldier efficiently isolates a part of the nest), and a low local graph redundancy (blocking of a tunnel forces attacking ants to take long detours). The values chosen for λ and ξ cover the range of the estimated values (from 0 to 0.4 in steps of 0.01), while the number of nodes cover the original nest sizes (200, 400, 600 and 1000 nodes).

III. RESULTS

A. Model validation

Figure 2 shows the degree distributions of the real networks and the simulated networks generated by the BBP model, RGG model, and RP model for each nest. The real networks and the simulated networks show qualitatively similar behaviors, with peaks around $k = 1, 2$ and exponential type decay of the right tail. However, all of the nests have a peak at $k = 1$ while the BBP model creates a peak at $k = 2$ for seven nests: the BBP model seems to create less dead ends than there exist in the real networks (see also Fig. 3). Nest P67 shows the worst fit with a larger tail in the model networks: note that this nest looks like two nest parts that have fused together during growth (see SM Fig. 6), a process not considered in the simulated networks. The degree distributions generated from the RGG model and the RP model show similar patterns as the distributions from the BBP model.

The largest connected component (LCC) of the network is fundamental for internal transportation since communication is not possible between disconnected components. Figure 4(a) represents the sizes of the largest connected components in real networks and simulated networks. The BBP model generated networks having comparable sizes of the largest connected components with the ones in real nests for most cases while the RGG model failed to generate the largest connected components of proper size in most cases. However, the performance of the RP model is comparable with the BBP model. It is notable that the discrepancy between real networks and networks generated by RP models is larger when the average degree is low. The blind cutting of random pruning can in this case increasingly affect important links and thus reduce LCC, while with higher average degrees many 'back-up' links exist that help preserve the LCC. We can observe this pattern also in the sensitivity analysis of the RP model (see Fig. 13 in [29]): with increasing probability to prune an edge (ξ) LCC quickly degrades, while it is better preserved in the BBP model (see Fig. 5 below). This indicates that betweenness

based pruning is a better strategy than random pruning to keep the largest connected component.

Figure 4(b) compares the average topological distance between any two nodes in the largest connected component, Fig. 4(c) the backbone link ratios, and Fig. 4(d) the local graph redundancy. The RGG statistics are not shown because computing them makes little sense when the LCC is too small. Overall both the BBP and RP model successfully reproduced the average distances, backbone link ratios, and local graph redundancies of the real networks. However, in the case of average distance neither BBP nor RP catch the variation between nests: Kendall's correlation coefficient τ with the real nests ($\tau_{BBP} = 0.30$ and $\tau_{RP} = 0.363$) is not significantly different from 0. In the case of LCC ($\tau_{BBP} = 0.939$ and $\tau_{RP} = 0.818$), backbone link ratio ($\tau = 0.606$ for both models) or local graph redundancy ($\tau_{BBP} = 0.636$ and $\tau_{RP} = 0.606$) τ is significantly different from 0. Given the plotted standard deviations notable differences can only be detected for nests C9, P67, P78, and T29 concerning average distance, nests C10 and P67 concerning backbone link ratio, and C10 and P67 concerning local graph redundancy. Again, nest P67 stands out in this comparison. See [29] for the full distributions of distances as in Fig. 2.

To compare the performances of the BBP model and the RP model quantitatively we define a Z -score for each metric X such that:

$$Z(X) = \frac{X_{real} - \langle X \rangle_{Model}}{\sigma_{Model}} \quad (5)$$

where $\langle X \rangle_{Model}$ is the average of X for the given model and σ_{Model} is the standard deviation. We show the Z -scores for each metric and each nest in Tables I-IV in [29]. Interestingly we found that the BBP model has a tendency to perform better than the RP model for *Cubitermes* and *Thoracotermes* nests while for the nests of *Proculitermes* the RP model performs better than the BBP model.

B. Calibrated parameters

For each nest (except P67 and P79b) we found parameter sets (λ_{BBP}, ξ_{BBP}) with $\epsilon < 0.025$ for the BBP models and (λ_{RP}, ξ_{RP}) with $\epsilon < 0.027$ for the RP model. Nest P67 with $\epsilon_{BBP} = 0.285$ and $\epsilon_{RP} = 0.283$ is again an exception. In the case of Nest P79b, the BBP model ($\epsilon_{BBP} = 0.125$) and the RP model ($\epsilon_{RP} = 0.129$) provide us the taller nests than real ones since the errors ϵ are mainly from the height of the models (i.e., from ϵ_1 in Eq. 2). Table I summarizes the principal nest properties and the estimated parameter sets.

Model parameter ξ does not depend on taxonomy (analysis of variance at the genus level, $F = 0.78$, $df = (2, 9)$, $p = 0.49$ for BBP and $F = 0.60$, $df = (2, 9)$, $p = 0.57$ for RP), but λ depends on Genus ($F = 6.61$, $df = (2, 9)$, $p = 0.017$ for BBP and $F = 6.16$, $df = (2, 9)$, $p =$

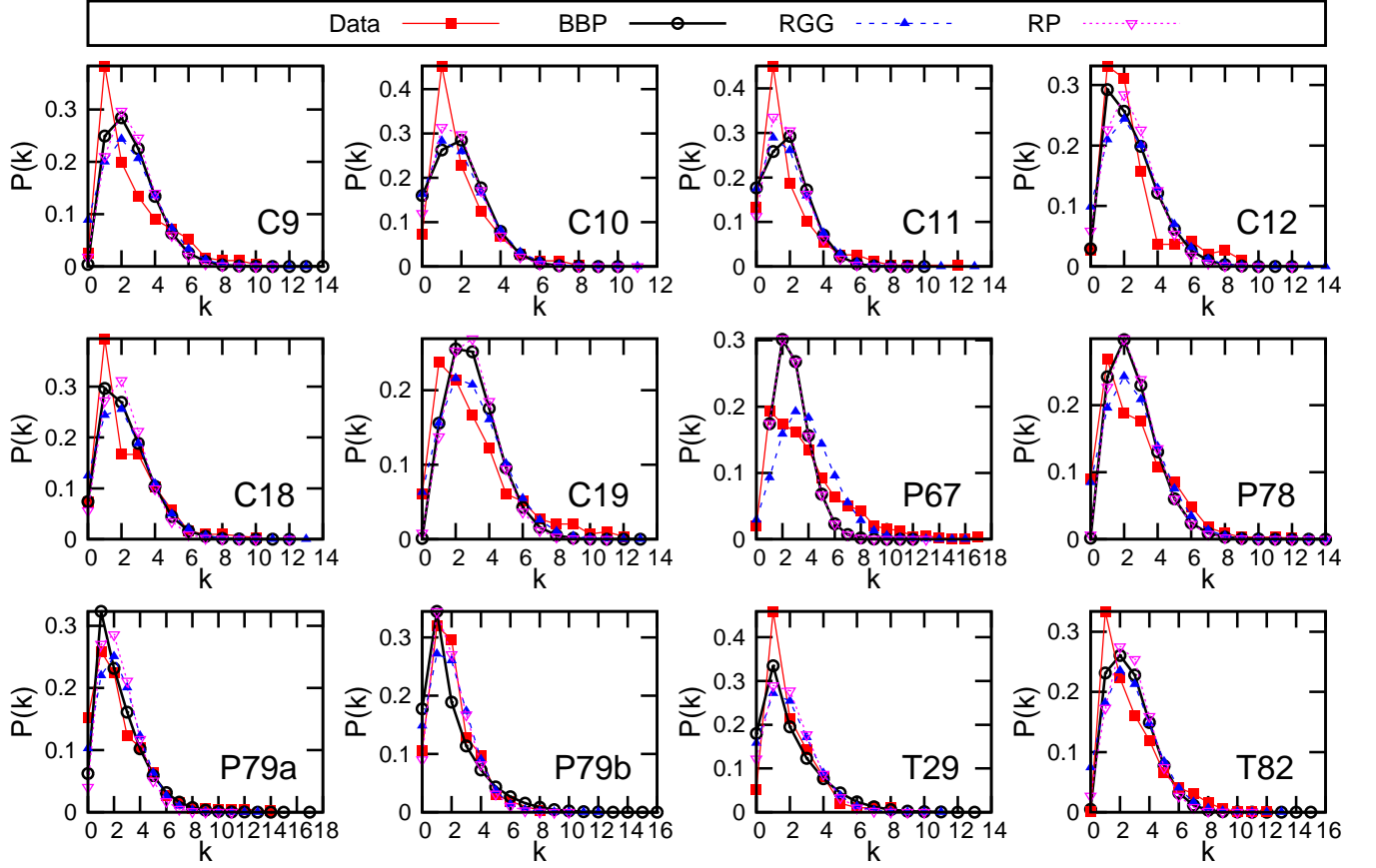


FIG. 2. (Color online) Degree distribution of the real networks (Data: red line) and the networks generated by the BBP model (BBP: black line), by the RGG model (RGG: blue line), and RP model (RP: magenta line). $P(k)$ are relative frequencies (summing to 1).

0.021 for RP; Tukey post-hoc: *Cubitermes* has significantly higher λ than *Procupitermes* and *Thoracotermes*).

C. Sensitivity to parameters (λ, ξ)

Figure 5 shows the sensitivity of network properties to the two free parameters (λ_{BBP}, ξ_{BBP}). We can observe that for the given parameter space, the size of the largest connected component is preserved when $\xi < 0.3$. We further see that average distance only depends on the decay rate λ , with low distances for low values of λ . Note that both sensitivities are quite different in the BP model (Fig. 13 in [29]): the network quickly breaks down with increasing ξ , especially for large λ , leading to small LCC's and low average distances in these LCC's. The nests have actually all values of $\lambda < 0.1$, indicating a moderate preference to continue construction from the most recent nodes. Backbone link ratio and local graph redundancy have a more complex dependence on (λ, ξ) (isolines seem to be linked to the product $\lambda\xi$), but termites could obtain a high backbone link ratio and a low local graph redundancy by increasing both λ and ξ .

All these qualitative observations are independent of nest size (number of nodes). The fact that the analyzed nests have nevertheless $\lambda < 0.1$ indicates that low average distances have more importance than increasing backbone link ratio or decreasing local graph redundancy. Only increasing ξ would further optimize the last three criteria, but also produce more disconnected chambers (decrease LCC) which incurs a cost to the colony (construct living space that cannot be used).

IV. DISCUSSION

In this paper we develop a simple network growth model (the BBP model) to test how the nest architecture of termites in the Termitinae sub-family emerges from self-organization based on local rules only. These nests consist of spherical chambers connected by tunnels, an architecture that can be represented as a network. The BBP model uses empirical nest descriptions (edge length and orientation) and two free parameters that control peripheral growth and pruning of existing edges. The free parameters are calibrated by fitting the model to the ob-

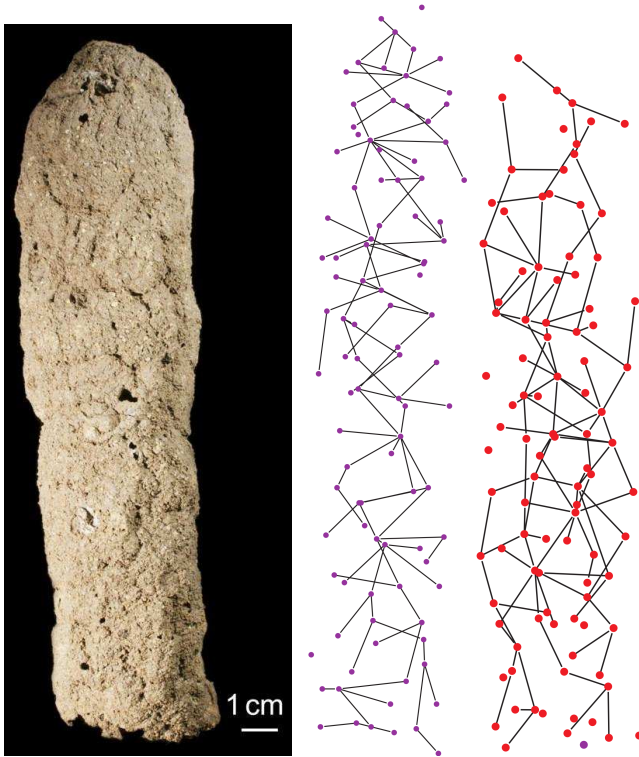


FIG. 3. (Color online) Analyzed *Thoracotermes* nest T29 with its network representation (middle) and a simulated network (right) by the BBP model. The pink node (lower right) indicates the initial node of the simulated network. Note that while there are isolated chambers in the original network, isolated chambers are more prevalent in the simulated nests. See the Discussion for further comments.

served nest height and the number of edges.

The BBP model correctly reproduces several emergent properties: a) node degree distribution (Fig. 2), b) the size of the largest connected component (Fig. 4(a)), c) the average distance between any two nodes in the largest connected component (Fig. 4(b)), d) the backbone link ratio (Fig. 4(c)), and e) the local graph redundancy (Fig. 4(d)). The properties (c) to (e) are of ecological relevance: the populations move around in these nests to place brood or larvae in chambers with optimal climatic conditions or to store and retrieve food (dead organic matter), short distances are therefore useful. It has already been shown for *Cubitermes* nests that their average distances are shorter than what could be obtained by randomly connecting the existing nodes [36], and our model correctly reproduces these lengths for most analyzed nests. The nest must also protect the colony against ant predators: this is done by the soldiers who can block a tunnel with their head capsule. A treelike network structure (or high backbone link ratio combined with local graph redundancy) helps with this strategy, and our model also correctly predicts these properties.

We compare our BBP model to two alternatives, the

RP model that is identical to BBP but with pruning applied to randomly chosen edges instead of the smallest betweenness centrality edge as in BBP, and to the Random Geometric Graph (RGG) model with the same space constraints as the two previous models. The latter can be ruled out because it cannot reproduce the observed largest connected component (LCC). The RP model generally performs as well as the BBP model (in the case of *Procubitermes* even better according to the Z-scores) for nests with average degree ≥ 2.0 . However, in nests with lower average degree (C10, C11, P79b, T29) RP tends to give fragmented networks. Furthermore, the sensitivity analysis shows that LCC size is conserved in the BBP model for a much larger range of pruning probabilities than in the RP model. Overall, pruning based on betweenness centrality (BBP) reproduces networks similar to the real networks more robustly than the RP model, we therefore conclude that betweenness centrality based pruning might be an important mechanism in termite nest construction.

While our model is intended to specifically reproduce the general properties and appearance of termite nests of the Termitinae sub-family, we can speculate that the key ingredients that regulate network morphogenesis in our model (peripheral growth and pruning) are likely to be shared also by a large number of biological transportation networks. For instance, pruning phenomena are observed in the maturation of neural networks (e.g. through programmed cell death [37] and synaptic pruning [38]). Mycelial networks formed by fungi similarly undergo a maturation process that involves peripheral growth of the hyphal tips in response to local changes of turgor pressure [39] and regression of filaments from nutrient depleted regions [40]. Growth of peripheral filaments that are subsequently pruned is also central to the formation of the network of cytoplasmic tubes that constitute the body of the plasmodium of *Physarum polycephalum* [41]. It seems plausible that self-organized network construction and optimization requires these mechanisms to operate on an initially highly connected network that is subsequently pruned. This could respond for instance to the fact that self-organized mechanisms can only evaluate transportation performance through the transportation itself, and the fine-tuning and optimization of the network requires to operate on already formed connections.

The sensitivity analysis (Fig. 5) indicates that our two free parameters λ (for peripheral growth) and ξ (for pruning) strongly influence both average distance and backbone link ratio (as well as local graph redundancy, another measure of how efficiently the nest can be defended, see [36]). We calibrated them by concentrating on a minimal number of features, nest height and total number of tunnels (i.e., links), in order to use the other nest features as emergent properties for model validation. The good qualitative agreement between these emergent features and the original nest features indicates that the estimated parameter values would not change much if

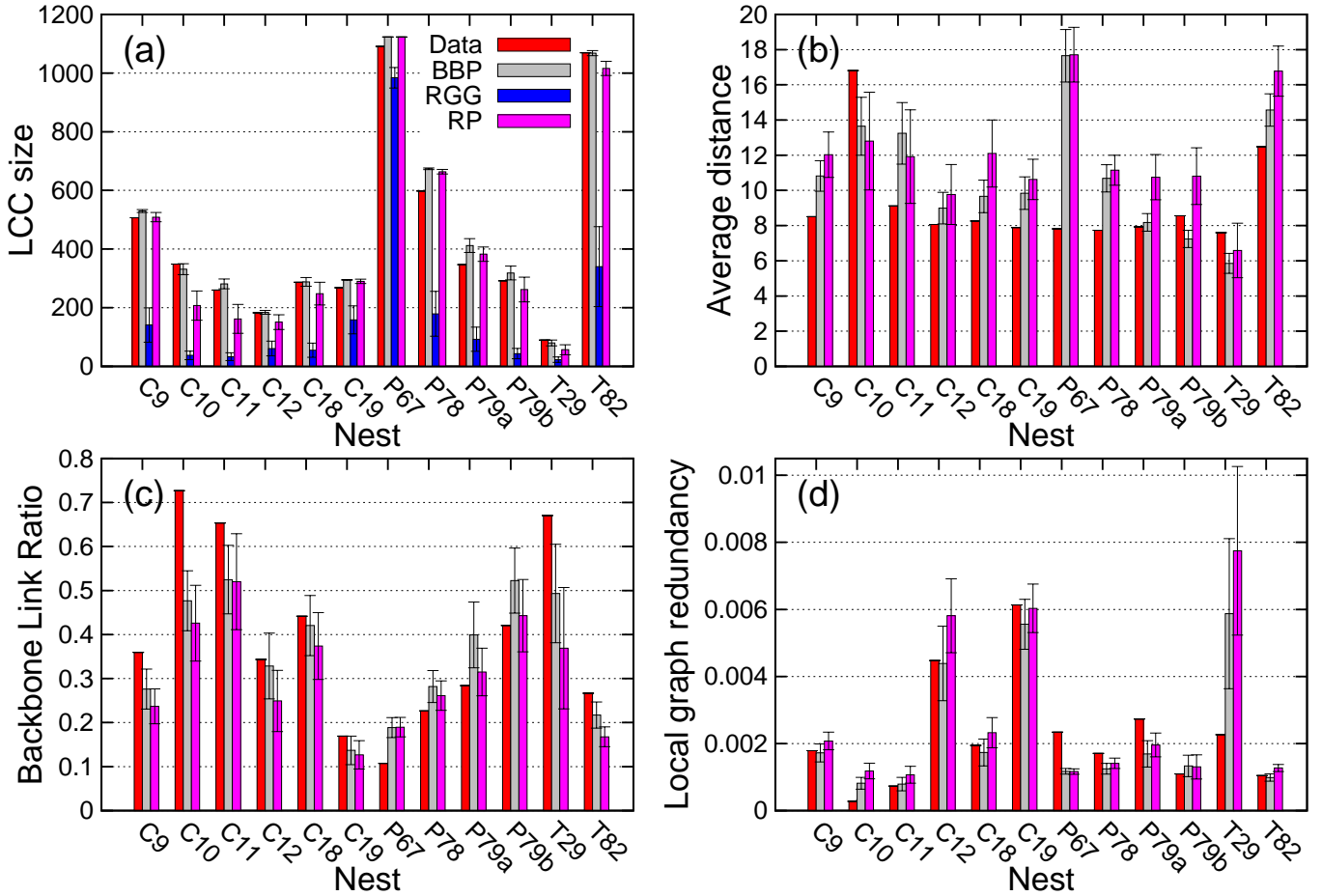


FIG. 4. (Color online) Comparison between the real networks (Data: red bar) and the networks generated by the BBP model (BBP: grey bar), by the RGG model (RGG: blue bar), and RP model (RP: magenta bar). (a) Size of the largest connected components (LCC). (b) Average topological distance between any two nodes in the LCC. (c) The backbone link ratio. (d) Local graph redundancy. The error bars represent standard deviations computed from 1000 model generated networks.

calibration had been based on more features. The peripheral growth parameter λ_{BBP} has mean value 0.033, meaning that the probability for a given node to be chosen for the next edge is divided by half every 30 newly added chambers. Typical *Cubitermes* nests grow in 1-3 months [30], meaning that the half life is of the order of some days to a week, a plausible duration for chemical marking in termites. In larger nests our model will also predict that growth becomes spotted on the nest surface, in agreement with field observations by one of the authors (CJ, unpublished).

Our model does not explain how termites decide to dig a new tunnel or how they choose an orientation in space - it is an empirical model at an intermediate scale. As such it resembles the 2D ant tunneling network models suggested by [5, 12, 42], or 2D termite tunneling models as suggested by [6–8]. However, though 3D network data of social insect nests become increasingly available [43–45], our model seems to be the first to predict the

nest’s 3D network architecture. Further work should investigate how nests grow in time [45] and, on the other end, how termites decide where and when to construct.

Another feature not predicted by our model are the characteristic “hats” atop the *Cubitermes* nests (see SM [29]). It is not known how termites decide to start constructing laterally, and we could not identify statistical properties specific to the height where the hat is constructed. The behavioral algorithm underlying this specific “hat” feature is therefore an open question.

There were also some misfits in this work. For example, nest P67 is badly explained. This nest actually consists of two columns that have grown together (see SM [29]). This bad fit might be corrected by letting the model grow from two randomly chosen initial nodes. Also, the above ground nest is only part of a termite colony, it connects to an extensive underground tunneling network through which termites access food. The absence of this underground network in our data leads to an underestimation

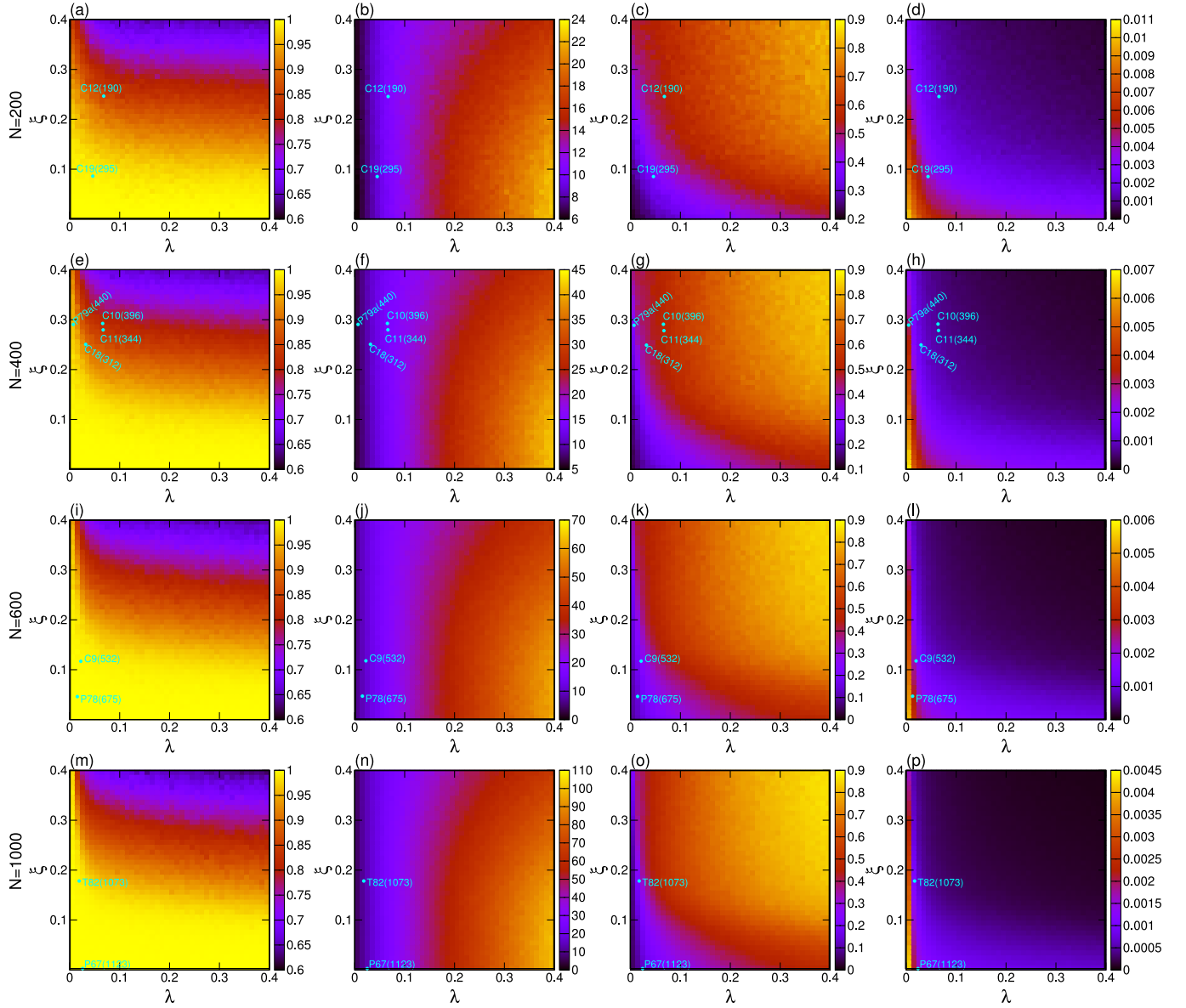


FIG. 5. (Color online) Sensitivity analysis of the BBP model. The sizes of largest connected components (LCC) are depicted in (a), (e), (i), and (m). LCC is shown as the fraction of the original network size. The average distances are depicted in (b), (f), (j), and (n). The backbone link ratios are depicted in (c), (g), (k), and (o). The local graph redundancies are depicted in (d), (h), (l), and (p). Note that nest P79b with $\xi = 0.451$ and nest T29 with $\xi = 0.551$ lie outside the range of simulated ξ -values, they therefore do not appear in figures (e) to (h) and (a) to (d) respectively. The spatial information of Nest C9 was considered as physical constraints for the analysis.

of edge-betweenness of the lower edges, thus explaining the often observed pruning of these edges (Fig. 3) that leads to an illogical disconnection between the above and below ground nest parts. This could only be corrected if one collects, in addition to the above ground part, a cast of the corresponding underground network [44]. Note also that underground tunnels are often built at a constant distance below ground, thus effectively leading to a 2D below ground network [25, 46]: an extended dataset could therefore explore how 3D networks connect to 2D networks and how this alters network properties. A further discrepancy is observed in the node degree distributions (Fig. 2): all nests have a peak at degree 1, but for half the nests the model predicts a peak at degree 2. This could probably be corrected by including this criterion when fitting λ , or by treating the fusion distance R (model part iv) as a free parameter. In the interest of keeping the model simple we refrained from such extended fitting procedures.

The sensitivity analysis (Fig. 5) shows that average distance only depends on the peripheral growth parameter λ , not on the pruning parameter ξ . Backbone link ratio and local graph redundancy, on the other side, depend on both parameters (higher λ can be compensated by lower ξ or the other way around). Both observations are true for all analyzed nest sizes. An efficient termite nest should have small average distances, high backbone link ratios and low local graph redundancy. The detected nest positions suggest that short distances are more im-

portant for the colony than the other two criteria (that can be linked to nest defense). However, further information is required on what termites actually do in their nests and on their vulnerability to predation before further speculating about such issues.

In sum, we found a parsimonious empirical network growth model based on self-organized principles that successfully predicts the nest architecture of Termitinae nests. Peripheral growth (i.e. some volatile chemical marking of new chambers/nodes/edges) and pruning of less important edges are important ingredients in this model.

ACKNOWLEDGMENTS

Y.-H. Eom acknowledges support of the EC FET project SIMPOL (No. 610704) and MULTIPLEX (No. 317532). We thank A Robert (IRD, Bondy) for *Thora-cotermes* nest T82, A Nel (National Museum of Natural History, Paris) for some *Cubitermes* nests and nest T29, and P Annoyer (SANGHA) for the other *Cubitermes* nests. We also thank the hospitals of Toulouse (Pr Rousseau) and Tours (Pr D Herbreteau) for scanning our nests. Finally, the authors thank Dr A. Akpesse for helping to collect the *Proculitermes* nests in Côte d'Ivoire. This work was supported by the French ANR grant ANR-06-BYOS-0008.

-
- [1] A. Perna and T. Latty, *Journal of the Royal Society Interface* **11**, 20140334 (2014)
 - [2] J. Buhl, K. Hicks, E. R. Miller, S. Persey, O. Alinvi, and D. J. T. Sumpter, *Behavioral Ecology and Sociobiology* **63**, 451 (2009)
 - [3] A. Perna, B. Granovskiy, S. Garnier, S. C. Nicolis, M. Labédan, G. Theraulaz, V. Fourcassie, and D. J. T. Sumpter, *PLoS Computational Biology* **8**, e1002592 (2012)
 - [4] D. Griffon, C. Andara, and K. Jaffe, *Emergence, self-organization and network efficiency in gigantic termite-nest-networks build using simple rules*, Tech. Rep. arXiv:1506.01487 (arXiv, 2015) <http://arxiv.org/pdf/1506.01487v1.pdf>
 - [5] J. Buhl, J. Gautrais, J. Louis Deneubourg, P. Kuntz, and G. Theraulaz, *J Theor Biol* **243**, 287 (Dec 2006)
 - [6] S.-H. Lee, P. Bardunias, and N.-Y. Su, *Journal of Theoretical Biology* **243**, 493 (2006)
 - [7] I. Haifig, C. Jost, V. Janei, and A. M. Costa-Leonardo, *Animal Behaviour* **82**, 1409 (2011)
 - [8] C. Jost, I. Haifig, C. R. R. de Camargo-Dietrich, and A. M. Costa-Leonardo, *Insectes Sociaux* **59**, 369 (2012)
 - [9] S. Camazine, J.-L. Deneubourg, N. R. Franks, J. Sneyd, G. Theraulaz, and E. Bonabeau, *Self-organization in biological systems* (Princeton University Press, Princeton, 2001) p. 538
 - [10] D. Levinson and B. Yerra, *Transportation Science* **40**, 179 (2006)
 - [11] N. R. Franks, N. Gomez, S. Goss, and J.-L. Deneubourg, *Journal of Insect Behavior* **4**, 583 (1991)
 - [12] J. Gautrais, J. Buhl, S. Valverde, P. Kuntz, and G. Theraulaz, *PLoS One* **9**, e109436 (2014)
 - [13] E. Strano, V. Nicosia, V. Latora, S. Porta, and M. Barthélemy, *Scientific reports* **2** (2012)
 - [14] A. Tero, R. Kobayashi, and T. Nakagaki, *Information and Material Flows in Complex Networks Information and Material Flows in Complex Networks*, *Physica A: Statistical Mechanics and its Applications* **363**, 115 (4 2006), <http://www.sciencedirect.com/science/article>
 - [15] Q. Ma, A. Johansson, A. Tero, T. Nakagaki, and D. J. T. Sumpter, *Journal of the Royal Society Interface* **10**, 20120864 (2013)
 - [16] S. Valverde, B. Corominas-Murtra, A. Perna, P. Kuntz, G. Theraulaz, and R. V. Sole, *Physical Review E* **79**, 066106 (2009)
 - [17] M. Barthélemy and A. Flammini, *Journal of Statistical Mechanics: Theory and Experiment* **2006**, L07002 (2006), <http://stacks.iop.org/1742-5468/2006/i=07>
 - [18] M. Barthélemy and A. Flammini, *Physical Review Letters* **100**, 138702 (2008)
 - [19] S. Itzkovitz and U. Alon, *Physical Review E* **71**, 026117 (Part 2 FEB 2005), <http://dx.doi.org/10.1103/PhysRevE.71.026117>

- [20] R. Albert and A.-L. Barabási, Reviews of modern physics **74**, 47 (2002)
- [21] M. E. J. Newman, SIAM Rev **45**, 167 (2003), <http://arxiv.org/pdf/cond-mat/0303516.pdf>
- [22] L. d. F. Costa, O. N. Oliveira Jr, and G. Travieso, Advances in Physics **60**, 329 (2011)
- [23] M. Barthélemy, Physics Reports **499**, 1 (2011), <http://arxiv.org/abs/1010.0302v2>
- [24] F. Xie and D. Levinson, Modeling the Growth of Transportation Networks: A Comprehensive Review **9**, 291 (2009)
- [25] P. Jmhasly and R. H. Leuthold, Insectes Sociaux **46**, 332 (1999)
- [26] C. Noirot and J. Darlington, in *Termites: evolution, sociality, symbioses, ecology*, edited by T. Abe, D. E. Bignell, and M. Higashi (Kluwer Academic Publishers Dordrecht, 2000) Chap. 6, pp. 121–139
- [27] W. R. Tschinkel, Journal of Insect Science **10**, 1 (2010)
- [28] A. Perna, C. Jost, E. Couturier, S. Valverde, S. Douady, and G. Theraulaz, Naturwissenschaften **95**, 877 (2008)
- [29] “See supplemental material at <http://link.aps.org/supplemental/10.1103/PhysRevE.92.062410> for a detailed statistical description of each nest,”
- [30] C. Noirot, in *Biology of Termites vol II*, edited by K. Krishna and F. M. Weesner (Academic Press, 1970) pp. 73–125
- [31] G. Josens and K. Soki, Insectes Sociaux **57**, 303 (2010)
- [32] O. H. Bruinsma, *An analysis of building behaviour of the termite Macrotermes subhyalinus*, Ph.D. thesis, Lanbouwhoogeschool te Wageningen, Belgium (1979)
- [33] R. Jander and K. Daumer, Insectes Sociaux **21**, 45 (1974)
- [34] M. Girvan and M. E. J. Newman, Proc Natl Acad Sci U S A **99**, 7821 (Jun 2002)
- [35] P. Bardunias and N.-Y. Su, Animal Behaviour **78**, 755 (2009)
- [36] A. Perna, S. Valverde, J. Gautrais, C. Jost, R. V. Solé, P. Kuntz, and G. Theraulaz, Physica A **387**, 6235 (2008)
- [37] R. W. Oppenheim, Annual review of neuroscience **14**, 453 (1991)
- [38] G. Chechik, I. Meilijson, and E. Ruppín, *Neural Computation* **10**, 1759 (Oct. 1998), ISSN 0899-7667, <http://dx.doi.org/10.1162/089976698300017124>
- [39] R. R. Lew, Nat Rev Micro **9**, 509 (Jul. 2011), ISSN 1740-1526, <http://dx.doi.org/10.1038/nrmicro2591>
- [40] R. E. Falconer, J. L. Bown, N. A. White, and J. W. Crawford, Proceedings of the Royal Society B: Biological Sciences **272**, 1727 (2005)
- [41] D. Kessler, in *Cell Biology of Physarum and Didymium*, Vol. 1, edited by C. Aldrich and J. W. Daniel (Academic Press, New York, 1982) pp. 145–208
- [42] E. Toffin, D. Di Paolo, A. Campo, C. Detrain, and J.-Deneubourg, Proc Natl Acad Sci U S A **106**, 18616 (Nov 2009)
- [43] W. R. Tschinkel, J Insect Sci **10**, 88 (2010)
- [44] W. R. Tschinkel, J Insect Sci **11**, 105 (2011)
- [45] N. J. Minter, N. R. Franks, and K. A. R. Brown, J R Soc Interface **9**, 586 (Mar 2012)
- [46] A. A. Moreira, L. C. Forti, A. P. P. Andrade, M. A. C. Boaretto, and J. F. S. Lopes, Studies on Neotropical Fauna and Environment **39**, 109 (2004)